PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
H04N 5/14, 7/46

A1

(11) International Publication Number: WO 95/21503

(43) International Publication Date: 10 August 1995 (10.08.95)

(21) International Application Number: PCT/GB95/00207

(22) International Filing Date: 2 February 1995 (02.02.95)

(30) Priority Data:
9401989.0 2 February 1994 (02.02.94) GB

(71) Applicant *(for all designated States except US)*: BRITISH BROADCASTING CORPORATION [GB/GB]; Broadcasting House, London W1A 1AA (GB).

(72) Inventor; and
(75) Inventor/Applicant *(for US only)*: SARGINSON, Peter [GB/GB]; British Broadcasting Corporation, Broadcasting House, London W1A 1AA (GB).

(74) Agent: ROBSON, Aidan, John; Reddie & Grose, 16 Theobalds Road, London WC1X 8PL (GB).

(81) Designated States: GB, JP, US, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.

P102798RR

5

(54) Title: VIDEO MOTION ESTIMATION

(57) Abstract

A method is provided for estimating motion vectors providing image displacements between reference images (IN) in a sequence of digital video images and intermediate image (BN) falling between reference images in sequence. Motion vectors are derived to estimate a prediction (PN) of one reference image (IN) from another in the sequence. Motion vectors are then derived for generating predictions of each intermediate image (BN) between two reference images (IN) in the sequence from one of the two reference images (IN) or their predictions (PN). From these estimated vectors further vectors are derived for deriving a prediction of each intermediate image (BN) from the other one of the two reference images or their predictions which it falls between.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	GB	United Kingdom	MR	Mauritania
AU	Australia	GE	Georgia	MW	Malawi
BB	Barbados	GN	Guinea	NE	Niger
BE	Belgium	GR	Greece	NL	Netherlands
BF	Burkina Faso	HU	Hungary	NO	Norway
BG	Bulgaria	IE	Ireland	NZ	New Zealand
BJ	Benin	IT	Italy	PL	Poland
BR	Brazil	JP	Japan	PT	Portugal
BY	Belarus	KE	Kenya	RO	Romania
CA	Canada	KG	Kyrgystan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo			SE	Sweden
CH	Switzerland	KR	Republic of Korea	SI	Slovenia
CI	Côte d'Ivoire	KZ	Kazakhstan	SK	Slovakia
CM	Cameroon	LI	Liechtenstein	SN	Senegal
CN	China	LK	Sri Lanka	TD	Chad
CS	Czechoslovakia	LU	Luxembourg	TG	Togo
CZ	Czech Republic	LV	Latvia	TJ	Tajikistan
DE	Germany	MC	Monaco	TT	Trinidad and Tobago
DK	Denmark	MD	Republic of Moldova	UA	Ukraine
ES	Spain	MG	Madagascar	US	United States of America
FI	Finland	ML	Mali	UZ	Uzbekistan
FR	France	MN	Mongolia	VN	Viet Nam
GA	Gabon				

VIDEO MOTION ESTIMATION

This invention relates to video motion estimation of the type which can be used in a video coder for reducing the bit rate of a digital video signal such as may be required for storage on digital storage media or for broadcast.

Various proposals for reducing the bit rate of a digital video signal have been made such as the MPEG-1 and MPEG-2 algorithms from the Motion Pictures Expert Group (MPEG). MPEG-1 is now an international standard and is specified in ISO/IEC document IS 11172 parts 1,2 and 3. A description of the MPEG-1 video compression algorithm can be found in Communications of the ACM April 1991, Vol.34, No. 4 MPEG-2 is a substantial extension of MPEG-1 and is due for publication as an international standard in 1995.

Both MPEG algorithms define the pictures comprising a video sequence as being one of three types. These are as follows:

A. Intra pictures (I) are coded without reference to other pictures. They serve as access points to the coded video sequence where decoding can begin.

B. Predicted pictures (P) are coded with reference to a motion compensated prediction derived from a previous I or P picture. Coding of P pictures is more efficient than for I pictures.

C. Bi-directionally predicted pictures (B) are coded with reference to a forward motion compensated prediction from a previous I or P picture and a backward motion compensated prediction from a future I or P picture and provide the greatest degree of compression.

The order of the three picture types within a video sequence is not constrained and will generally depend on the requirements of the application.

A typical sequence of I, P and B pictures is illustrated in Figure 1. Here, frame I1 is an I picture, frames P4 and P7 are P pictures and frames B2, B3, B5, B6, B8 and B9 are B pictures. Frame P4 is coded with reference to a forward prediction derived from frame I1. Frame P7 is coded with reference to a forward prediction derived from frame P4. Frames B2 and B3 are coded with reference to a forward prediction derived from frame I1 and a backward prediction derived from frame P4. Frames B5 and B6 are coded with reference to a forward prediction derived from frame P4 and a backward prediction derived from frame P7. Frames B8 and B9 are coded with reference to a forward prediction derived from frame P7 and a backward prediction derived from frame I10.

When a video coder is implemented in hardware, a large part of its circuitry is devoted to the measurement of motion vectors in order to generate the required motion compensated predictions. In the case of B pictures, two simultaneous motion measurements are required in order to generate both the forward and backward predictions. In the case of the MPEG-2 algorithm, twice as many motion measurements are required as both 'frame' and 'field' versions of each motion vector are needed.

We have appreciated that the number of motion measurements, and hence the amount of motion measurement circuitry, may be considerably reduced if some additional processing of the forward motion vectors is performed. Such processing may be implemented in either hardware or software.

The invention is defined in the appended claims to which reference should now be made.

Specific embodiments of the invention will now be described in detail by way of example with reference to the accompanying drawings in which:

Figure 1 shows the sequence of I, B, and P frames described above;

Figure 2 shows a conventional motion estimator arrangement required to produce vectors defining the sequence of frames shown in Figure 1; and

Figure 3 shows a motion estimator embodying the invention.

An MPEG picture sequence of the type shown in Figure 1, which is in display order and where I, P and B refer to the MPEG picture type is as follows.

B6, B7, P8, B9, B10, P11, B12, B13, P14, B15, B16, I17, B18, B19, P20......etc

The sequence may be split into "triples" of {BBP} or (less often) {BBI}, i.e. two interpolated and one reference picture.

For each {BBP} triple, a coder's motion estimator must generate five motion vectors. For example, in the case of the triple {B12, B13, P14} the five required motion vectors are:

P11 to P14	(Forward prediction of P14 from P11)
P11 to B12	(Forward prediction of B12 from P11)
P11 to B13	(Forward prediction of B13 from P11)
P14 to B12	(Backward prediction of B12 from P14)
P14 to B13	(Backward prediction of B13 from P14)

For each {BBI} triple a coder's motion estimator must generate four motion vectors. For example, in the case of the triple {B15, B16, I17} the four required motion vectors are:

P14 to B15	(Forward prediction of B15 from P14)
P14 to B16	(Forward prediction of B16 from P14)
I17 to B15	(Backward prediction of B15 from I17)
I17 to B16	(Backward prediction of B16 from I17)

To date, the architecture which is generally proposed to generate these motion vectors is based around two motion estimators as shown in Figure 2. This comprises a forward motion estimator 2 receiving a previous reference picture and the current picture and producing forward motion vectors and a backward motion vector estimator 4 receiving a future reference picture

and the current picture and producing backward motion vectors. These are the vectors referred to above.

Table 1 shown below describes the passage of the sequence, B12, B13, P14, B15, B16, I17, through the motion estimators. It will be seen that the sequence of the pictures is re-ordered to P14, B12, B13, I17, B15, B16, prior to input to the motion estimators in accordance with the MPEG algorithm. Each column of the table represents an instant in time. For example, the first column means, "when P14 is on the current picture input, P11 is applied to the previous reference picture input and the forward motion vector output generates the motion vector P11 to P14". The table demonstrates how the two motion estimators of Figure 1 can together generate the required five motion vectors for each {BBP} triple and the required four motion vectors for each {BBI} triple.

Table 1

Current Picture Input	P14	B12	B13	I17	B15	B16
Previous Reference Picture Input	P11	P11	P11	P14	P14	P14
Future Reference Picture Input		P14	P14		I17	I17
Forward Motion Vector Output	P11 to P14	P11 to B12	P11 to B13		P14 to B15	P14 to B16
Backward Motion Vector Output		P14 to B12	P14 to B13		I17 to B15	I17 to B16

The amount of hardware required for the videc coder can be reduced if only a single one of the motion

estimators of Figure 2 is used. This becomes feasible if further vector processing is used to generate the additional motion vectors. An arrangement for a single motion vector estimator is shown in Figure 3.Table 2 illustrates the passage of the re-ordered sequence P14, B12, B13, I17, B15, B16 through the motion estimator of Figure 3.

Table 2

Current Picture Input	P14	B12	B13	I17	B15	B16
Reference Picture Input	P11	P11	P11	P14	P14	P14
Motion Vector Generated	P11 to P14	P11 to B12	P11 to B13	P14 to I17	P14 to B15	P14 to B16

The required motion vectors for the {BBP} triple are (P11 to P14), (P11 to B12), (P11 to B13), (P14 to B12) and (P14 to B13). The motion estimator has generated the first three of these. The missing two motion vectors may be calculated using additional vector processing as follows:

(P14 to B12) = (P11 to B12) - (P11 to P14)
(P14 to B13) = (P11 to B13) - (P11 to P14)

Similarly, the required motion vectors for the (BBI) triple are (P14 to B15), (P14 to B16), (I17 to B15) and (I17 to B16). The motion estimator has generated the first two of these as well as the vector (P17 to I17). The missing two motions vectors may be calculated as follows:

(I17 to B15) = (P14 to B15) - (P14 to I17)
(I17 to B16) = (P14 to B16) - (P14 to I17)

As long as the forward motion estimator generates the correct forward motion vectors, then the derived backward motion vectors are correct. This holds for objects moving with linear and non-linear motion which represents the majority of picture material. the method can break down if the image contains repetitive patterning (e.g. a chess board). In such a case, the forward motion estimator may identify a number of motion vectors that will yield equally good predictions even though they need not represent the actual motion of the object. If the forward motion vector is incorrect, the derived backward motion vector may yield a poor prediction. (It should be noted that if a prediction resulting from a backward motion vector is poor, the coder will simply choose some other prediction mode for that particular macroblock and the erroneous backward motion vector will go unnoticed).

The additional vector processing required to generate all forward and backward vectors using a single vector estimator could be implemented in hardware using vector adders and/or subtractors and some suitable storage and switching arrangement to supply the vectors to the adders at the correct times. Alternatively the processing could be carried out in software.

The current MPEG-2 test model details a further improved bit-rate reduction standard comprising field and frame motion vectors. The idea is as follows.

MPEG-2 defines a macroblock as covering an area of picture 16 pixels by 16 frame lines, 8 lines originating from the odd field and the other 8 lines originating from the even field. The job of a coder's prediction generator stage is to generate "predictions" of the macroblocks comprising the current frame. If the image is moving then these predictions will generally be realised by motion compensating (i.e. assigning one or more motion vectors to each macroblock of a reference frame.

At present the motion estimators (of the type referred to above) generate both 'field' and 'frame' motion vectors in order to derive two predictions of each macroblock in the current frame:- one based on the field motion vectors, the other based on the frame motion

vector. The coder chooses the prediction which most closely matches the actual macroblock of the current frame.

In the case of frame motion vectors, a macroblock is treated as a single 16 x 16 block of a frame. A single "frame" motion vector is assigned to describe how a prediction of the macroblock may be generated by motion compensating a 16 x 16 pixel area of a reference frame.

In the case of field motion vectors two "field" motion vectors are used to independently generate predictions of the odd field and even field components of a macroblock. i.e. One field motion vector is used to predict the 8 odd lines of the macroblock from the odd field of a reference picture and the other field motion vector is used to predict the 8 even field lines from the even field of a reference picture.

To date, the proposals for implementing field and frame motion vectors in hardware make use of two independent motion estimator blocks, one performing frame motion vector generation, the other performing field motion vector generation. Both motion estimators represent the same amount of hardware. This is because, although the motion estimator concerned with field motion vector generation produces twice as many motion vectors as the other, it needs only to compare [8X16] blocks of image data as opposed to the [16X16] blocks processed by the frame motion estimator.

The motion vectors are used to generate two predictions of a macroblock from the current picture, one generated using the frame motion vector, the other generated using the two field motion vectors. It is then the job of the coder's "decision module" to decide which is the better prediction.

If the motion seen by the motion estimator in the odd field is different from that seen in the even field then the two field motion vectors will be different and the field motion vector based prediction is highly likely to be the best.

If the motion seen by the motion estimator is the same in both fields then the two field motion vectors will

be identical and further, they will probably match the frame motion vector. In this case there is unlikely to be any significant difference between the two predictions.

We propose that the frame motion estimator can be dispensed with. A frame motion vector may be derived instead by averaging the two field motion vectors. If the two field motion vectors match, then the frame motion vector will have the same value resulting in similar frame and field based predictions. however, if the two field motion vectors are different, the frame motion vector is likely to be incorrect and the field based prediction is likely to be best. The vector generation is explained in more detail below.

Motion occurring in the odd and even fields of a previous frame is independently measured and used to generate predictions of the add and even fields of a future frame, i.e. the odd field of the future frame is predicted from the odd field of the previous frame and correspondingly with the even field. This one prediction of a single future frame comprising odd and even fields has been generated.

Next the vectors measured in the odd field are averaged with those in the even field to give a set of frame motion vectors. A further prediction of the future frame may now be generated by motion compensating the previous frame with the set of frame motion vectors.

Thus there are two predictions of a single specific future frame. One is derived by independently measuring the motion occurring in each field of a previous frame and motion compensating each field accordingly. The other is derived by averaging the motion vectors from the two fields of a previous frame to yield a set of frame motion vectors and using these to compensate the whole of the previous frame.

This combination of field vectors can be implemented in dedicated hardware or in software in the motion estimator.

CLAIMS

1. A method for estimating motion vectors defining image displacements between reference images in a sequence of digital video images and intermediate images falling between reference images in the sequence the method comprising the steps of estimating motion vectors for deriving a prediction of one reference image from another in the sequence, estimating motion vectors for deriving a prediction of each intermediate image between two reference images in the sequence from one of the two reference images or their predictions, and from the estimated vectors deriving further vectors for deriving a prediction of each intermediate image from the other one of the said two reference images or their predictions which it falls between.

2. A method according to claim 1 in which the thus derived vectors conform with a predetermined bit-rate reduced video signal standard.

3. A method according to claim 1 in which each image consists of two or more fields and where motion occurring in each field is measured independently in order to generate a set of field motion vectors for each field and the sets of field motion vectors are combined to produce a single set of frame motion vectors, thereby enabling two predictions of a future image to be made.

4. A method for video motion estimation between images in a sequence in which each image consists of two or more fields and where motion occurring in each field is measured independently in order to generate a set of field motion vectors for each field and the sets of field motion vectors are combined to produce a set of frame motion vectors thereby enabling two predictions of a future image to be made.

5. A method for estimating motion vectors defining image displacements between reference frames in a sequence

of digital video images and intermediate frames falling between frames in the sequence substantially as herein described.

6. A method for video motion estimation substantially as herein described.

FIG. 1

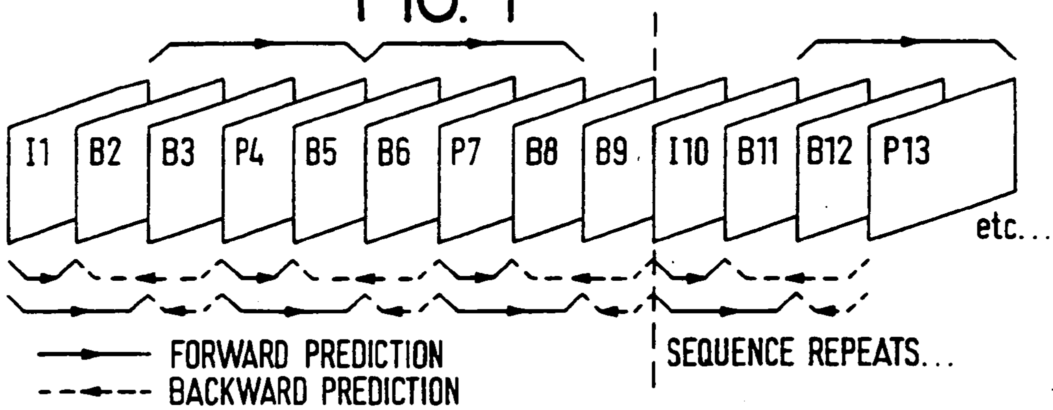


FIG. 2

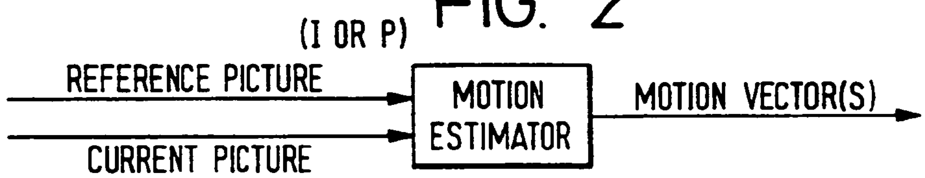
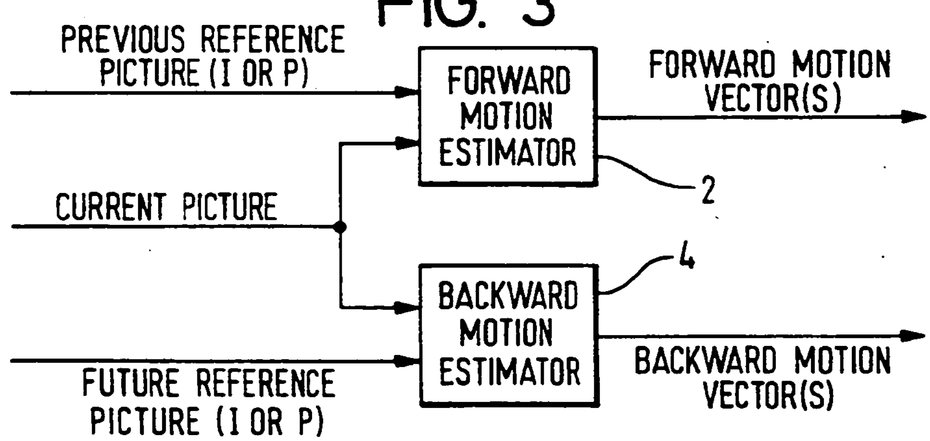


FIG. 3



INTERNATIONAL SEARCH REPORT

Inte[illegible]nal Application No
PCT/GB 95/00207

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 H04N5/14 H04N7/46

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 H04N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	EP,A,0 534 350 (MATSUSHITA) 31 March 1993 see the whole document ---	1,2,5,6
A	EP,A,0 541 389 (MATSUSHITA) 12 May 1993 see the whole document ---	3,4
A	US,A,5 049 991 (NIIHARA) 17 September 1991 see abstract ---	1
A	EP,A,0 424 026 (SONY) 24 April 1991 -----	

☐ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
29 March 1995	25.04.95

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Yvonnet, J

INTERNATIONAL SEARCH REPORT

Information on patent family members

Inter nal Application No
PCT/GB 95/00207

Patent document cited in search report	Publication date	Patent family member(s)		Publication date
EP-A-0534350	31-03-93	US-A-	5198901	30-03-93
		CA-A-	2078950	24-03-93
		JP-A-	6217293	05-08-94
EP-A-0541389	12-05-93	JP-A-	6030395	04-02-94
		JP-A-	5130594	25-05-93
		AU-B-	637289	20-05-93
		CA-A,C	2082280	09-05-93
		US-A-	5369449	29-11-94
US-A-5049991	17-09-91	JP-A-	2219391	31-08-90
EP-A-0424026	24-04-91	JP-A-	3129985	03-06-91
		JP-A-	3129984	03-06-91
		JP-A-	3129986	03-06-91
		AU-B-	6336294	21-07-94
		US-A-	5132792	21-07-92

Form PCT/ISA/210 (patent family annex) (July 1992)